

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-239913

(43)Date of publication of application : 12.09.1995

(51)Int.Cl.

G06K 9/62

G06K 9/03

G06K 9/66

(21)Application number : 06-029792

(71)Applicant : TOSHIBA CORP

(22)Date of filing : 28.02.1994

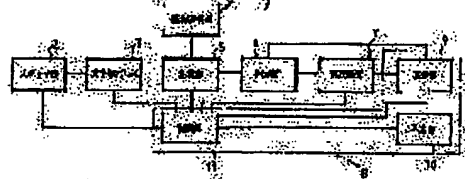
(72)Inventor : Kiyono Kazuji
Michiniwa Kenichi

(54) CHARACTER RECOGNITION DEVICE

(57)Abstract:

PURPOSE: To provide a character recognition device with excellent real time performance in which a mis-reading rate is reduced through learning.

CONSTITUTION: The character recognition device 1 has a recognition section (1st recognition means) 4, a discrimination section 6, a re-recognition section (2nd recognition means) 7, and a correction means 8. When the recognition section 5 decides a character pattern candidate with a highest similarity (evaluation value) in the reading stage, the judging section 6 judges the candidate. When the judging section 6 judges the candidate liable to incur mis-reading, the re-recognition section 7 applies recognition processing to an input character pattern corresponding to the candidate. A correction means 8 corrects a judging criterion and a processing criterion so that the judging section 6 and the re-recognition section 7 make surer judgement and recognition based on the judging result by the judging section 6 and the recognition result by the re-recognition section 7 in the learning stage.



LEGAL STATUS

[Date of request for examination] 10.09.1999

[Date of sending the examiner's decision of rejection] 08.01.2002

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

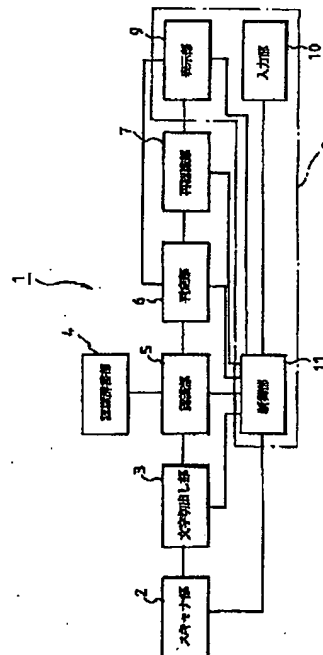
[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

BEST AVAILABLE COPY

(11)特許出願公開番号



【特許請求の範囲】

【請求項1】 入力文字パターンと認識辞書とを照合して評価値を求め、入力文字パターンの認識結果として最も評価値の高い文字パターン候補を決定する第1の認識手段と、

この第1の認識手段により決定された文字パターン候補が誤読を招き易い候補か否かを変更可能な判定基準に従って判定する判定手段と、

この判定手段により誤読を招き易い候補と判定された場合に、その候補に対応する入力文字パターンについて変更可能な処理基準に従って特徴抽出を行い、その文字を認識する第2の認識手段と、

前記判定手段の判定結果及び第2の認識手段の認識結果に基づいて前記判定手段及び第2の認識手段がそれぞれより確実な判定及び認識を行えるように前記判定基準及び処理基準を修正し得る修正手段と、
を有することを特徴とする文字認識装置。

【請求項2】 入力文字パターンと認識辞書とを照合して評価値を求め、入力文字パターンの認識結果として比較的评价値の高い複数の文字パターン候補を決定する第1の認識手段と、

この第1の認識手段により決定された最も評価値の高い文字パターン候補が誤読を招き易い候補か否かを第1の認識手段が求めた各評価値に基づき変更可能な判定基準に従って判定する第1の判定手段と、

この第1の判定手段により誤読を招き易い候補と判定された場合に、その候補に対応する入力文字パターンについて変更可能な処理基準に従って特徴抽出を行い、その文字を認識する第2の認識手段と、

この第2の認識手段により認識された文字が前記第1の認識手段により決定された複数の文字パターン候補のうち上位所定数の文字パターン候補に有るか無いかの判定を行い、無い場合はその文字をリジェクトする第2の判定手段と、

前記判定手段の判定結果及び第2の認識手段の認識結果に基づいて前記判定手段及び第2の認識手段がそれぞれより確実な判定及び認識を行えるように前記判定基準及び処理基準を修正し得る修正手段と、
を有することを特徴とする文字認識装置。

【請求項3】 少なくとも前記判定手段及び第2の認識手段をニューラルネットで構成したことを特徴とする請求項1又は2記載の文字認識装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、文字認識装置に関し、より詳しくはニューラルネットを用いた文字認識装置に関する。

【0002】

【従来の技術】 近年、ニューラルネットを用いて文字認識を行うことが試みられている。

【0003】 このニューラルネットを用いた文字認識装置は、ある入力文字パターンの認識結果に誤りが生じた場合に、その入力文字パターンに対する次回からの文字認識において正解の文字パターン候補が得られるような方向性を持ってニューラルネットワークにおけるアーク荷重値を修正することで、入力文字パターンの特徴の変化に応じた良好な文字認識を安定して行うことが可能とされている。

【0004】

【発明が解決しようとする課題】 しかしながら、ニューラルネットを用いた従来の文字認識装置は、全ての文字に対してニューラルネットによる文字認識を行っていたため、リアルタイム性に欠けるという問題があった。

【0005】 そこで、本発明は、上記事情に鑑みてなされたものであり、学習により誤読率の低減を図ることができ、しかもリアルタイム性に優れた文字認識装置を提供することを目的とするものである。

【0006】

【課題を解決するための手段】 上記目的を達成するために請求項1記載の文字認識装置は、入力文字パターンと認識辞書とを照合して評価値を求め、入力文字パターンの認識結果として最も評価値の高い文字パターン候補を決定する第1の認識手段と、この第1の認識手段により決定された文字パターン候補が誤読を招き易い候補か否かを変更可能な判定基準に従って判定する判定手段と、この判定手段により誤読を招き易い候補と判定された場合に、その候補に対応する入力文字パターンについて変更可能な処理基準に従って特徴抽出を行い、その文字を認識する第2の認識手段と、前記判定手段の判定結果及び第2の認識手段の認識結果に基づいて前記判定手段及び第2の認識手段がそれぞれより確実な判定及び認識を行えるように前記判定基準及び処理基準を修正し得る修正手段とを有することを特徴とするものである。

【0007】 また、請求項2記載の文字認識装置は、入力文字パターンと認識辞書とを照合して評価値を求め、入力文字パターンの認識結果として比較的评价値の高い複数の文字パターン候補を決定する第1の認識手段と、この第1の認識手段により決定された最も評価値の高い文字パターン候補が誤読を招き易い候補か否かを第1の認識手段が求めた各評価値に基づき変更可能な判定基準に従って判定する第1の判定手段と、この第1の判定手段により誤読を招き易い候補と判定された場合に、その候補に対応する入力文字パターンについて変更可能な処理基準に従って特徴抽出を行い、その文字を認識する第2の認識手段と、この第2の認識手段により認識された文字が前記第1の認識手段により決定された複数の文字パターン候補のうち上位所定数の文字パターン候補に有るか無いかの判定を行い、無い場合はその文字をリジェクトする第2の判定手段と、前記判定手段の判定結果及び第2の認識手段の認識結果に基づいて前記判定手段及

び第2の認識手段がそれぞれより確実な判定及び認識を行えるように前記判定基準及び処理基準を修正し得る修正手段とを有することを特徴とするものである。

【0008】また、請求項3記載の文字認識装置は、少なくとも前記判定手段及び第2の認識手段をニューラルネットで構成したことを特徴とするものである。

【0009】

【作用】請求項1記載の文字認識装置を読取段階と学習段階とに分けて説明する。

【0010】読取段階においては、第1の認識手段が最も評価値の高い文字パターン候補を決定すると、判定手段はその候補に対し判定を行う。そして、誤読を招き易い候補と判定された場合には、第2の認識手段はその候補に対応する入力文字パターンについて認識処理を行う。学習段階においては、判定手段の判定結果が誤っていた場合は、修正手段により判定手段がより確実な判定を行えるように判定基準を修正する。また、第2の認識手段の認識結果が誤っていた場合は、修正手段により第2の認識手段がより確実な認識を行えるように処理基準を修正する。判定手段の学習が進むに従い、誤読を招き易い候補か否かの判定がより確実となり、第2の認識手段の学習が進むに従い、第2の認識手段による認識がより確実となり、誤読率が低減する。

【0011】また、請求項2記載の文字認識装置によれば、第2の認識手段による再認識結果の正誤を第2の判定手段により判定しているので、誤読率がより低減する。

【0012】また、請求項3記載の文字認識装置によれば、修正手段により判定手段及び第2の認識手段を構成しているニューラルネットにおけるアーク荷重値を修正することで、判定基準及び処理基準を迅速に修正できる。

【0013】

【実施例】以下、本発明の実施例を図面を参照して詳述する。

【0014】図1は本発明の文字認識装置の一実施例を示す概略ブロック図である。

【0015】この実施例装置1は、文書のイメージを検出するスキャナ部2と、このスキャナ部2により検出された文書のイメージから1文字毎に文字パターンを切り出す文字切出し部3と、この文字切出し部3からの入力文字パターンと認識辞書部4に格納された文字パターン候補とを照合して複数の文字パターン候補を決定する第1の認識手段としての認識部5と、この認識部5により決定された文字パターン候補が誤読を招き易い候補か否かを変更可能な判定基準に従って判定する判定部6と、この判定部6により誤読を招き易い候補と判定された場合に、その候補に対応する入力文字パターンについて変更可能な処理基準に従って特徴抽出を行い、その文字を認識する第2の認識手段としての再認識部7と、判定部

6の判定結果及び再認識部7の認識結果に基づいて判定部6及び再認識部7がそれぞれより確実な判定及び認識を行えるように前記判定基準及び処理基準を修正し得る修正手段8とを有して概略構成されている。

【0016】次に、上記各部の詳細を説明する。

【0017】前記スキャナ部2は、文書上に光を照射する光源と、文書からの反射光を受けて電気信号に変換する光电変換素子とを具備し、文書全体を光学的に走査して文書のイメージを検出するものである。

【0018】前記認識辞書部4には、文字切出し部3からの入力文字パターンとの照合の対象となる多数の文字パターン候補が格納されている。

【0019】前記認識部5は、入力文字パターンと認識辞書部4に格納された文字パターン候補とを照合し、例えばパターンマッチング法により評価値として類似度を求め、入力文字パターンの認識結果として比較的類似度の高い複数の文字パターン候補を文字コードで決定し、類似度と共に判定部6に出力するものである。なお、評価値は、類似度に限らず、他の評価値でもよい。また、評価方法は、パターンマッチング法に限定されず、他の方法でもよい。

【0020】前記修正手段8は、例えばCRTディスプレイの如き表示部9と、キーボード、マウス等を備えた入力部10と、判定部6の判定基準及び再認識部7の処理基準を修正すると共にこの装置1の各部を制御する制御部11とを具備している。

【0021】前記判定部6は、例えば図2に示すようなニューラルネットにより構成されている。すなわち、この判定部6は、複数の入力ノード60aからなる入力層60と、各入力ノード60aとアーク63を介してネットワークで結合された複数の中間ノード61aからなる中間層61と、各中間ノード61aとアーク64を介してネットワークで結合された一対の出力ノード62aからなる出力層62とから構成されている。

【0022】各入力ノード60aには、認識部5によるパターンマッチング法で算出された各文字パターン候補に対する類似度（第1位乃至第n位）がそれぞれ入力されるようになっている。出力ノード62aからは、判定結果が制御部11に出力されるようになっている。

【0023】また、各アーク63、64には、各ノード60a、61a、62a間の結合の強さを示す荷重値（実数）が設定されている。この荷重値の設定により、判定部6が判定を行うための判定基準が設定されたことになる。

【0024】認識部5にて算出された各類似度が、各入力ノード60aに入力されると、設定された判定基準により、認識部5によって認識された最も類似度の高い文字パターン候補すなわち第1位の文字パターン候補が誤読を招き易い候補か否かが判定され、その判定結果が出力ノード62aから制御部11に出力されるようになっ

5

ている。例えば、誤読を招き易い候補と判定した場合は、「1」が出力され、その他の場合は「0」が出力される。

【0025】また、各荷重値は、表示部8に表示された認識結果のうち誤読文字にカーソルを合わせて正解の文字を入力する等の入力部10の操作に基づいて、制御部*

$$y = f(\sum \omega_i \cdot x_i)$$

で表される。一方、事例データとして $(y'(s), x_i'(s) | s=1, \dots, S)$ が存在しており、荷重値*

$$R(\omega) = \sum_{i=1}^S (y'(s) - y(\omega, x_i'(s)))^2 \quad \dots(2)$$

で表される。ここで、式(1)に $x_i = x_i'$ (判定部6が誤判定したときの類似度)の入力を与えたときに、教師データ $(\{1\}, y')$ に近い出力 $y (= f(\sum \omega_i \cdot x_i'))$ が得られるように、 ω_i を少しずつ変更して、誤差関数 $R(\omega)$ を最も小さくする ω_i を求める。これを各ノードについて行い、判定部6の処理基準を修正する。

【0026】また、判定基準としては、例えば、文字パターン候補の第1位の類似度が所定値より低い、又は文字パターン候補の第1位と第2位との類似度の点数差が所定値以下の場合は、誤読を招き易い候補と判定し、それ以外を正読と判定する。例えば、文字パターン候補の第1位、第2位、第3位及び第4位の類似度が、それぞれ100点、90点、80点、70点の場合は、第1位と第2位との類似度の点数差が小さいので、第1位の文字パターン候補を誤読を招き易い候補と判定する。また、文字パターン候補の第1位が100点、第2位以降の全てが0点の場合は、第1位と第2位との点数差が大

【0027】前記再認識部7は、例えば図3に示すようなニューラルネットにより構成されている。すなわち、この再認識部7は、複数の入力ノード70aからなる入力層70と、各入力ノード70aとアーク73を介してネットワークで結合された複数の中間ノード71aからなる中間層71と、各中間ノード71aとアーク74を介してネットワークで結合された複数の出力ノード72aからなる出力層72とから構成されている。

【0028】各入力ノード70aには、文字切出し部3にて切り出された文字パターンのうち判定部6にて誤読を招き易い候補と判定されたイメージの文字パターンを0、1のパターンで入力されるようになっている。出力ノード72aからは、認識結果が制御部11に出力されるようになっている。

【0029】また、各アーク73、74には、各ノード70a、71a、72a間の結合の強さを示す荷重値(実数)が設定されている。この荷重値の設定により、再認識部7が認識するための処理基準が設定されたことになる。

6

*11により例えば、誤差逆伝播(バックプロパゲーション)法に従って修正されるようになっている。これらの荷重値を修正することにより、判定基準が学習され修正されることになる。すなわち、ノードが1つの場合を例にして説明すると、そのノードの出力 y は、入力データを x_i 、荷重値を ω_i とすると、

$$\dots(1)$$

※ ω のノードに $x'(s)$ を入力した出力値を $y(\omega, x'(s))$ とすると、誤差関数 $R(\omega)$ は、

【0030】誤読を招き易い候補と判定された文字パターンが、入力ノード70aに入力されると、設定された処理基準により、特徴抽出が行われ、文字が文字コードとして認識され、その認識結果が出力ノード72aから制御部11に出力されるようになっている。

【0031】また、各荷重値は、表示部8に表示された認識結果のうち誤読文字にカーソルを合わせて正解の文字を入力する等の入力部10の操作に基づいて、制御部11により例えば、誤差逆伝播(バックプロパゲーション)法に従って修正されるようになっている。これらの荷重値を修正することにより、処理基準が学習され修正されることになる。すなわち、判定部6と同様に、前記式(1)に $x_i = x_i'$ (再認識部7が誤読した文字パターン)の入力を与えたときに、教師データ(正しい文字コード) y' に近い出力 $y (= f(\sum \omega_i \cdot x_i'))$ が得られるように、 ω_i を少しずつ変更して、誤差関数 $R(\omega)$ を最も小さくする ω_i を求め、これを各ノードについて行い、再認識部7の処理基準を修正する。

【0032】なお、前記判定部6及び再認識部7を構成するニューラルネットの学習法は、ここではバックプロパゲーション法とするが、これに限定されない。また、層数は上記3層に限定されず、4層以上でもよい。また、判定部6及び再認識部7は、学習可能なものなら、ニューラルネットに限定されず、他の構成としてもよい。

【0033】次に、本実施例の動作を図6のフローチャートに従い読取段階と学習段階とに分けて説明する。

【0034】まず、読取段階において、スキャナ部2が、制御部11の制御の下に、文書全体を光学的に走査して文書のイメージを検出すると、文字切出し部3は、文書のイメージを1文字毎に切り出して、認識部5に出力する。ここで、その切り出された文字パターンのうちの1つが図4に示すように「閉」とする。

【0035】認識部5は、その文字パターンについて認識辞書部4に格納された文字パターン候補を用いて認識処理を行い、類似度の高い順に所定数の文字パターン候補及び類似度を判定部6に出力する(S1)。その出力例を図5に示す。同図に示すように、文字パターン候補

及び類似度は、それぞれ第1位に「関」（240点）、第2位に「閉」（222点）、第3位に「開」（203点）、第6位に「門」（158点）が出力されたとする。

【0036】判定部6は、第1位の文字パターン候補（図5の例では「関」）が誤読を招き易い候補か否かの判定を行い、その判定結果を制御部11に出力する（S2）。例えば、誤読を招き易い候補と判定した場合は、それを示す「1」を出力し、正読となる候補と判定した場合は、それを示す「0」を出力する。

【0037】制御部11は、判定部6の判定結果を解析し（S3）、判定結果が「1」（誤読を招き易い候補）と判定された場合には、それに対応する文字パターン（図4の例では「閉」）を再認識部7に入力する。再認識部7は、制御部11の制御の下に、その文字パターン（「閉」）について処理基準に従って特徴抽出を行い、その文字を認識する（S4）。

【0038】その認識結果は、制御部11の制御の下に、表示部8に表示される（S5）。この場合は、「関」が表示され、誤読となっているとする。

【0039】次に、学習段階においては、操作者は、表示部8に表示された認識結果と読取対象の文書とを見比べて、誤読文字「関」を発見すると、入力部10を操作して、カーソルを誤読文字「関」に合わせ正解の文字「閉」を入力する（S6）。

【0040】続いて、前記ステップS3における判定部6の判定結果を解析し（S7）、その判定結果が「0」となっていた場合は、認識部5で誤読され、判定部6で誤判定された結果とみなせるので、制御部11は、図5に示すような各文字パターン候補に対する類似度を判定部6に入力し、出力が「1」（誤読を招き易い候補）となるように判定部6における判定基準を修正して学習を行う（S8）。これにより、より確実な判定が行えるようになる。

【0041】前記ステップS7にて、判定結果が「1」となっていた場合は、認識部5で誤読され、判定部6で正しく判定され、更に再認識部7で誤読された結果とみなせるので、制御部11は、文字パターン「閉」を再認識部7に入力し、出力が「閉」となるように再認識部7における処理基準を修正して学習を行う（S9）。これにより、より確実な認識が行えるようになる。

【0042】このような本実施例によれば、判定部6の学習が進むに従い、誤読を招き易い候補が否かの判定が*

*より確実となり、再認識部7の学習が進むに従い、再認識部7による認識がより確実となり、誤読率の低減を図ることができる。また、修正手段8により判定部6及び再認識部7を構成しているニューラルネットにおけるアーク荷重値を修正することで、判定基準及び処理基準を迅速に修正でき、リアルタイム性に優れたものとなる。

【0043】本実施例の他の実施例としては、再認識部7により認識された文字が認識部5により決定された複数の文字パターン候補のうち上位所定数の文字パターン候補に有るか無いかの判定を行い、無い場合はその文字をリジェクトする第2の判定手段を設けてもよい。これにより、誤読率の低減をより図ることができる。

【0044】なお、本発明は上記実施例に限定されず、その要旨を変更しない範囲内で種々に変形実施できる。

【0045】

【発明の効果】以上詳述した請求項1記載の発明によれば、判定手段の学習が進むに従い、誤読を招き易い候補か否かの判定がより確実となり、第2の認識手段の学習が進むに従い、第2の認識手段による認識がより確実となるので、誤読率の低減を図ることができる。

【0046】また、請求項2記載の発明によれば、第2の認識手段による再認識結果の正誤を第2の判定手段により判定しているので、誤読率の低減をより図ることができる。

【0047】また、請求項3記載の発明によれば、判定手段及び第2の認識手段を構成しているニューラルネットにおけるアーク荷重値を修正することで、判定基準及び処理基準を迅速に修正できるので、リアルタイム性に優れたものとなる。

【図面の簡単な説明】

【図1】本発明の光学的文字読取装置の一実施例を示す概略ブロック図。

【図2】本実施例の判定部の構成図。

【図3】本実施例の再認識部の構成図。

【図4】本実施例の入力文字パターンの一例を示す図。

【図5】本実施例の認識部の出力例を示す図。

【図6】本実施例の動作を示すフローチャート。

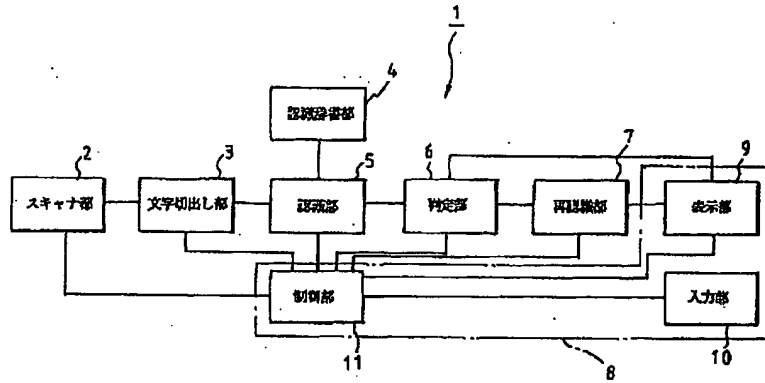
【符号の説明】

- 1 文字認識装置
- 5 認識部（第1の認識手段）
- 6 判定部
- 7 再認識部（第2の認識手段）
- 8 修正手段

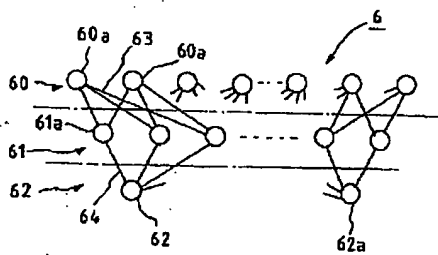
【図4】

閉

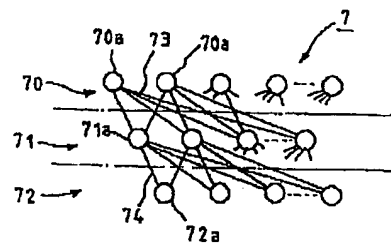
【図1】



【図2】



【図3】



【図5】

	文字パターン振幅	類似度
第1位	開	2 4 0
第2位	閉	2 2 2
第3位	開	2 0 3
...
第6位	閉	1 5 8
...

【図6】

